DOCUMENT RESUME

ED 349 817 FL 020 636

AUTHOR	French, Russell L.
TITLE	Portfolio Assessment and LEP Students.
PUB DATE	Aug 92
NOTE	37p.; In: Focus on Evaluation and Measurement. Volumes 1 and 2. Proceedings of the National Research Symposium on Limited English Proficient Student Issues (2nd, Washington, DC, September 4-6, 1991); see FL 020 630.
PUB TYPE	Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Information Analyses (070)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; *Limited English Speaking; *Portfolios (Background Materials); Program Evaluation; Scoring; *Student Evaluation; Teacher Education
IDENTIFIERS	*Portfolio Approach

ABSTRACT

This paper focuses on the use of portfolio assessment in the evaluation of limited-English-proficient (LEP) students. Specific sections address the following: recent developments in performance assessment; observations regarding LEP students and assessment; student portfolios (what they are and how they can be used); what a portfolio should contain; uses of portfolios in student, teacher, and program assessment; portfolio design issues; and scoring and rating portfolios of LEP students. Responses to the paper by Alice J. Kawakami and Daniel Koretz are appended. (VWL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Portfolio Assessment and LEP Students

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Russell L. French
University of Tennessee

ED349817

The Arguments For Alternative Forms of Assessment

There is both national and international demand for alternatives to present forms of student assessment. We find that demand expressed in the National Educational Goals, the products of the National Educational Goals panel and the "AMERICA 2000" strategy designed to flesh out those goals.

We also find it in publications and statements from various nationally prominent groups and in a number of state educational reform initiatives. For example, the National Governors' Association in its recent publication From Rhetoric to Action, states:

> There is considerable activity in new test development at the state and national levels by consortia of states and traditional test publishers. ...The goals are the same: creating instruments that go beyond paper-and-pencil, multiple-choice tests pegged to national norms to those that capture understanding and measure performance against high standards.[1]

In setting forth its nine-point educational agenda, Essential Components of A Successful Education System[2], the National Business Roundtable calls for a new education system which is performance- or outcome-based, and it states, "Assessment strategies must be as strong and as rich as the outcomes." The National Alliance For Restructuring Education and the National Center on Education and the Economy also demand a restructured education system that is performance-based. They state:

> A performance-based education system requires high standards and challenging goals for students, world-class curriculum and instruction that are demanding and varied, new performance assessments that measure higher-order skills, incentives for continuous improvement for students and educators, and consequences for persistent failure to improve.[3]

To understand this widespread interest in new assessments and assessment methodologies, one must understand the concerns about current testing programs. Arguments against current forms of assessment and for alternatives include:

FL 020 636

BEST COPY AVAILABLE

1. Current standards for student performance as reflected in our tests are not high enough to meet the needs of the next century (or even today). New standards are needed. New assessments suitable for all students are needed.

2. Current tests and student evaluation procedures do not measure what all students actually know and are able to do.

3. Current standardized tests do not measure what is taught; i.e., they are not aligned with most curricula.

4. Current tests and assessment procedures do not measure adequately the higher order thinking skills and processes needed in today's and tomorrow's world, skills in which students are demonstrating weakness. Alternative, authentic assessments are needed.

5. Curriculum must be built around real life (authentic) tasks. Only real life, authentic assessments can validly and adequately assess the results of such a curriculum.

6. New assessments that can be used to compare the educational progress of school systems, schools, and individual students both nationally and internationally over time are needed.

7. To be appropriate for all students, assessments must be criterion-referenced; i.e., they must measure gains in knowledge and skills over time.

I believe most of these arguments are self-explanatory to most readers from the educational community. That is not to say that most readers agree with all of them but, together, they form the basis for the demand for new assessment technologies and the resultant activity. Later in this paper, the relevance of some of these points to assessment by portfolio, even at a class-room level, should be apparent.

Recent Developments In Performance Assessment

A significant amount of experimentation in new or refined methods of performance assessment is in progress. Much of the effort focuses on authenticity or realism of the assessments (tests), the standards against which to measure student performance, procedures for rating or scoring the new assessments, and training of educators in how to use and score them. That work is under-way in individual states and school districts and in projects of national scope. Among the well-known "national" or multi-state projects are the New Standards Project directed by Dr. Lauren Resnick (University of Pittsburgh Learning Research and Development Center) and Dr. Marc

Tucker (National Center On Education And The Economy), the Coalition of Essential Schools headed by Dr. Ted Sizer, the State Alternative Assessment Exchange initiated by the Council of Chief State School officers, and the projects (e.g., Project Zero, Project Propel) being implemented collaboratively by Harvard University and several school districts in several states.

Individual states already engaged in development of alternative assessments to replace current standardized tests include Arizona, California, Connecticut, Kentucky, Maryland and Vermont. A great many other states are contemplating restructuring of their assessment programs to include performance assessments. Among those with policy or legislation in place or near acceptance are Alabama, Colorado, Georgia, Iowa, South Carolina, and Virginia. Perhaps the most far-reaching state effort currently is that in Kentucky, where legislation mandates that an entire new assessment program consisting of performance assessments and NAEP-like tests be in place within five years. Further, the assessments created are to be appropriate for all learners in the schools, and all teachers in the state are to be trained to score the assessments and to produce similar ones for use in their own instruction. A 28 million dollar contract for this work has just been let.

Clusters of states also are discussing establishment of consortia to develop new performance assessments, both as a means of offsetting high development costs and as a means of creating assessments with meaning beyond the boundaries of a single state. It is now clear that the movement of families from place to place, especially within a geographic region, requires that sound assessment data follow the student. To appropriately place and instruct students in restructured curricula and schools; administrators and teachers must know what each individual knows and is able to do. This "cluster" activity is supported by the urging of President Bush, Secretary of Education Alexander, the National Goals Panel and the National Council on Standards and Testing to create new American Achievement Tests "capable of comparing the performance of students both nationally and internationally." It is thought that these American Achievement Tests should not be a single set of assessments developed at a national level but sets of assessments developed by clusters of states with similar curriculum frameworks and educational situations. "Cluster assessments" can then be equated to each other to provide national norms.

Three key emphases in these state, regional and national initiatives should be noted by classroom teachers and those responsible for instructing and measuring the progress of LEP students. First, there is great concern that new assessments be valid and appropriate for all students, regardless of handicap or language. Second, it is understood that if performance assessments are to replace current stan-

dardized tests, the methodologies used in those assessments must also be used in ongoing instruction. As has always been true, assessments must be aligned with what is taught and how it is taught if assessment results are to be valid. Third, there is concern that assessment be a part of instruction not apart from instruction. Therefore, there is emphasis on training teachers and administrators to develop and use performance assessments.

Observations Regarding LEP Students and Assessment

I do not pretend to be an expert in the education of limited and/or non-English proficient students. Indeed, my personal experience with these students is extremely limited. However, careful reading of recent literature on LEP learners and their instruction, discussions with persons responsible for teaching these students, experience in developing performance assessment instruments for both educators and students, recent experience with the RJR Nabisco Foundation's Next Century Schools (many have LEP students), teaching experience with at risk, K-12 learners, and some degree of common sense combine to lead me to several points for consideration by those who must assess the academic progress and ultimate achievements of LEP students.

1. <u>There is obviously a need to assess what LEP students really know and are able to do</u>. At issue in any assessment are its validity and reliability. In their simplest form, these concepts represent two questions: "How do I know that what I am measuring is what I really wanted to measure? (validity) How do I know that I am measuring consistently? (reliability)? Those issues are no less important to classroom assessments developed by teachers than they are to standardized tests. Experience in standardized testing has taught us that the language skills of the test taker influence his or her performance on the test, even when that test taker is supposedly English proficient. When a test is influenced in that way, the test is invalid for that particular learner. The invalidity stems from the fact that the measurement becomes a measurement of language rather than a measurement of whatever else we wanted to measure.

2. <u>There appears to be a need to reinforce a student's native language, not destroy it</u>. Several recent articles and papers on the instruction of LEP students report that the LEP student's self-concept, family relationships, and academic achievement suffer when instruction attempts to make him/her monolingual in English rather than bilingual or multi-lingual. Common sense also should tell us that we need an increasing number of persons pro-

ficient in two or more languages in our society to meet the increasing demands for international interaction. Why should we deplete or destroy some of our best resources?

If, then, we attempt to reinforce a student's native language in our instruction, we cannot do less in our assessments. Evaluation which allows only for the use of the English language sends a message quite contradictory to that being portrayed through instruction, and the "louder" message will be that sent through assessment. Whether we like it or not, assessment drives curriculum, or, more specifically, assessment drives students' perceptions of what is important in the curriculum. Further, assessment procedures inconsistent with instructional procedures also create an invalid test.

3. Learning styles and nonverbal communication patterns are critical to both instruction and assessment. This writer has researched the roles of both learning styles and nonverbal communication in the classroom for more than twenty years. For purposes of this paper, suffice it to say that there are at least seven different perceptual learning styles, to say nothing of varying cognitive, emotional and social styles. We know that there are learners who are print-oriented (dependent on reading), aural (dependent on listening), interactive (dependent on talking/verbalizing), visual (dependent on pictorial representations), haptic (dependent on touch and feel), kinesthetic (dependent on movement) and olfactory (dependent on smell and taste).[4] Further, greater numbers of certain types of learners are found in some cultures and backgrounds than in others.

Much also has been written about the importance of nonverbal communication in language and culture. More than 70 percent (perhaps as much as 90 percent) of whatever is communicated is communicated nonverbally.[5] Further, nonverbal cues do not have universal meaning. They carry different meanings in different cultures.[6] Much of language then is nonverbal, and many thought processes contributing to language are nonverbal.

It follows that instruction and assessment that do not take these differences among students, any students and especially LEP students, into account are likely to be unreliable and invalid much of the time. Learning style and nonverbal language influence language and assessment responses.

4. The two language systems possessed by bilingual students limit the value of assessment methods used currently. In her article in the ERIC/CUE Digest, Carol Ascher[7] concludes that individuals who are bilingual have two distinct but overlapping language systems that they rely on in different ways depending upon the situations in which they find themselves. Because of this phe-

nomenon, she is particularly concerned that "diagnostic protocols" for bilingual students include information beyond standardized test scores and that assessments more directly aligned with curriculum be developed. Ascher's points are important. If bilingual students change language systems with the situations and stresses that confront them, we can never be sure which language system has interpreted (or misinterpreted) the multiple choice test item and produced the response which we are scoring. Assessments that enable us to know what language system is at work are needed.

Performance assessments, particularly portfolio assessments, have much to offer in assessment of LEP students. Potentially, they can contribute much more knowledge than we now are obtaining about what these students really know and are able to do. They offer potentially greater validity and reliability than present testing technologies. Many portfolio entries can be done in the native language, thereby reinforcing bilinguality and accommodating language system shifts. Since portfolio entries need not be restricted to print, these assessments can accommodate differences in learning styles and nonverbal communication. However, none of these possibilities can become realities unless those desiring to use portfolio assessments understand (a) what student portfolios are, (b) how they can be used, and (c) how to design them.

Student Portfolios: What are they and how can they be used?

What Is a Portfolio?

Current work in developing performance assessments focuses on three assessment protocols or types: portfolios, performance tasks and exhibitions. Figure 1 provides definitions of each assessment type and a few key issues in their development and use. Careful study of the definitions in Figure 1 should enable the reader to see where and how these three types of performance assessment might overlap.

A portfolio might contain a number of performance tasks or assessments of those tasks. In many cases, performance tasks require construction, creation, description (written or oral), or other formats for task completion that lend themselves to portfolio inclusion. Often, performance tasks are quite structured in time and space. In a recent joint proposal with Educational Testing Service to develop performance assessments, we defined a performance task as any reality-based task which would require an hour and a half or less to complete.

An exhibition could include presentation of a portfolio of work, although that need not be the case. Or, a portfolio might contain assessments and photographs or other documentation of an exhibition. Obviously, an exhibition is a display of what has been produced over time. The emphasis is on display or presentation.

The reader should also be aware that performance assessment can take forms other than portfolios, performance tasks, or exhibitions. Instrumentation used in the performance evaluations of teachers and administrators historically has included observation records, interview protocols and self-reports. Similar forms of assessment can be used in evaluation of student learning and could be included in portfolios. In addition, there is considerable effort at this time to use computer-simulated tasks as substitutes for "real" tasks which often require substantial equipment and/or materials for each student being assessed. It appears that the same tasks transferred to computerized formats are more efficient and cost-effective while losing little or nothing in their validity, reliability, credibility, or effectiveness.

How Can Portfolios Be Used?

While the definition of a portfolio provided in Figure 1 is French's definition, it is very close to the definitions of others active in portfolio development and utilization. Dennie Palmer Wolf,[8] a research associate with Harvard's Project Zero, defines a portfolio as "a chronologically sequenced collection of work that records the evolution of artistic thinking." Paulson, Paulson and Meyer[9] define it as "a purposeful collection of student work that exhibits the student's efforts, progress, and achievements in one or more areas." These definitions immediately suggest certain attributes of a portfolio that may be helpful in assessing student progress. Note that they emphasize a collection of work(s), chronological organization and purposeful construction (i.e., construction with a goal or purpose).

Additional concepts important to the formulation of portfolio structures and uses are offered by several researchers and developers. Howard Gardner,[10] director of Project Zero, suggests that portfolios can best be used to assess a student's ability to produce, perceive, and reflect. Wolf[11] describes portfolios as contributing "biographies of work" (e.g., a biography of the development of a musical performance), ranges of works (e.g., a collection of diverse pieces) and reflections (student analyses of what they have produced). Resnick[12]

Figure 1
Performance Assessments

I. Portfolio

A purposeful, Chronological collection of student work, designed to reflect student development in one or more areas over time and student outcomes at one or more designated points in time.

Key Issues:

- assessment targets/exemplars/performance standards
- guidelines for inclusions
- scoring/rating procedures
- training of faculty

II. Performance Tasks

A reality-based task which can be completed within the confines of a single day or less.

Key Issues:

- realism of the task
- scoring/rating procedures
- performance standards

III. Exhibition

The presentation of a body of work which has taken place over several weeks, months, or years.

Key Issues:

- realism of the task(s)
- scoring/rating procedures
- performance standards
- guidelines for development

NOTE: Defenses or reflections by the student(s) are often used in combination with all three of the above listed assessments.

compares portfolio development and assessment to scouting in that students use the same process as "accumulating badges over a period of years," i.e., they will complete tasks and submit projects that they wish to use to demonstrate competence against published criteria. Many teachers and others involved in examination of student portfolios mention the insights produced by portfolio entries about student learning, both what is learned and how it is learned. These comments often focus on student demonstration of communication skills, psychomotor skills, artistic skills and thinking skills as well as knowledge acquisition. Learning process dimensions discussed include critical thinking, socialization, perseverance, self-criticism, on-time task completion, problem-solving strategies, pursuit of quality or high standards and student ability to pose and address meaningful questions.

Clearly, there exists in the current literature the notion that a portfolio has the capacity to, and should, produce a portrait of both learning outcomes and learning processes, a portrait that enables the viewer (assessor) to see what the producer (student) is capable of doing and how he/she thinks, works, develops. Assessment potential is both formative and summative.

Some of the criticism of current and historic student assessment practices is also useful in determining what the role of portfolios in future assessment models might be. In his recent Phi Delta Kappan article, Stiggins[13] bemoans the state of "assessment illiteracy" among American educators. He defines "assessment literates" as persons who can recognize that assessment targets are unclear, that assessment methods are missing their targets, that samples of performance are inadequate, that there are specific extraneous factors creeping into assessment data and that assessment results are unclear. He calls for programs to train educators at all levels to be "assessment literates," thereby enabling them to create new forms of student assessment which are more valid, reliable, and appropriate.

Wiggins,[14] like Stiggins, expresses a concern for the identification of clear assessment targets. However, he refers to those needed targets as standards which he defines as "educative, specific examples of excellence on tasks we value." Current student assessments lack these "concrete" benchmarks (or exemplars) for judging student work at essential tasks, Wiggins posits.

In this context, the measurement of student progress toward the exemplar or standard requires a series of successive approximations. In other words, what's missing in both large scale and local student assessments are clear, specifications of exit level results against which student work is continuously compared.

The assessment model being promoted is criterion-referenced rather than normative, longitudinal rather than periodic, and output rather than input driven. This reliance on output, particularly exit level outcomes, implies that student work might take several varied forms to which a common set of standards (criteria) must be applied. Common standards can be applied only to completed products, tasks, or performances, Wiggins argues.

There are sufficient implications for assessment by portfolio in the Wiggins and Stiggins articles to round out our conceptualization of the role of this device. Notice that a portfolio has the potential to display various stages of student progress toward a clearly defined standard/exemplar/assessment target if one is defined. It offers a longitudinal assessment method that can be closely matched to the assessment target. It offers a means of collecting multiple samples of diverse kinds of student work and results (products) which are concrete and usable in a variety of ways.

The focus of this discussion of portfolios has been their role in performance assessment, and that will continue to be the focus in the remainder of this paper. However, it should be noted that portfolios are often used as instructional devices. In fact, one of the current problems in portfolio development and utilization is the tension between instruction and assessment that many classroom teachers seem to feel.

Although assessment should be aligned with instruction, and assessment results should be used to direct subsequent instruction, these two processes have different parameters and requirements. Good teachers have long used monitoring (informal assessment) of student activities to make immediate adjustments in student tasks and in their own instructional practices. Further, they often feel obligated to give immediate assistance to students struggling with a task. Neither practice is appropriate to summative assessment in which validity and reliability of measurement must be maintained. When the purpose of the task or exercise is to determine a student's accomplishment of a prescribed standard or to determine progress toward that standard, students must be allowed to complete and submit products or productions for scoring without additional assistance. After assessment is completed, reteaching, additional review or additional practice can take place. It appears that portfolios tend to blur the critical lines between instruction and assessment and between formative and summative assessment even more than present testing procedures for many educators.

What Should A Portfolio Contain?

There is no simple answer to this question. Obviously, the type of portfolio, its storage and retrieval system, the subject areas, skills,

and processes involved in the assessment and the characteristics of the student(s) have to be considered in determining type and number of portfolio entries. Currently, various portfolios include written materials (essays, stories, themes, compositions, research papers, etc.), anecdotal information (logs, journals), work samples (selected seatwork, homework), projects/products (things created by the student or representations of them), tests/test scores, teacher comments/analyses, self-analyses, audiotapes, videotapes, photographs, drawings, paintings, observational records, and checklists. Notice that some of these items are the products of student activity, and some are assessments of student activities. However, the elements of the definition of a portfolio should be kept in mind. It is not a random collection of whatever is available, but a chronological collection of artifacts carefully chosen to represent the student's achievement of specified objectives and/or progress toward them. The outcomes being measured may be acquisition of knowledge, cognitive, psychomotor or social skills or attitudes and dispositions.

While the system for storing and retrieving information in a portfolio plays a significant role in determining types of entries, the limitations of space, time, and format are swiftly being erased by the technology now available. Linda Vista Elementary School in San Diego, California, has been experimenting for two years with a computerized portfolio that allows for computer storage and retrieval of multiple types of information including print, videotape, voice prints, and photographs. It is also interesting to note that Linda Vista Elementary School, an RJR Nabisco Foundation Next Century School, has more than 60 percent LEP students representing six native languages: Hispanic, Vietnamese, Cambodian, English, Laotian and Hmong. Students in the school are not grouped by age or grade level but by English proficiency, and aspects of the curriculum are taught in each native language. Obviously, the electronic portfolio is perceived as a means of accommodating a range of learners and languages and gathering data for assessment which transcends the boundaries of current standardized tests.[15]

Uses of Portfolios in Student, Teacher and Program Assessment

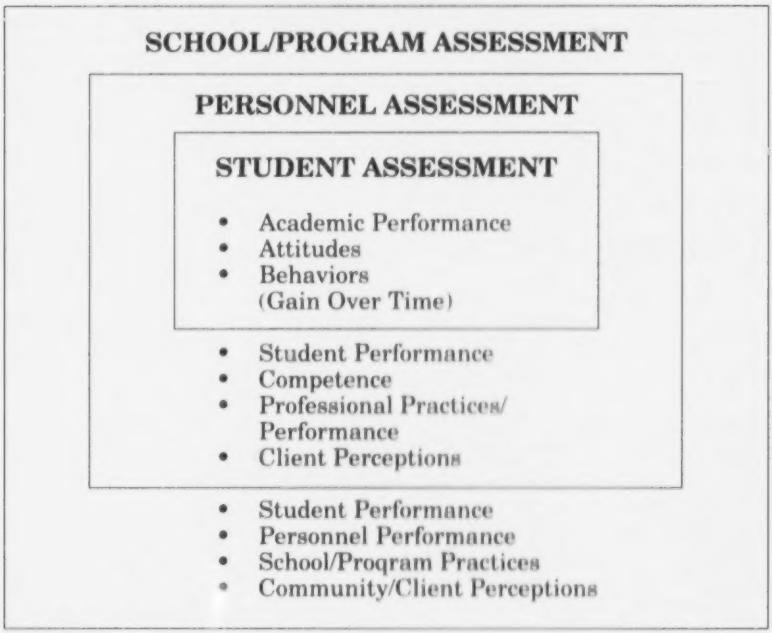
Figure 2 presents a detailed summary of a nested assessment model. At the center is student assessment; at the second level is personnel assessment, and at the third broadest level is school or program assessment. Each assessment level builds upon those encompassed within it. Clearly, student assessment is or should be at the core of the model, and portfolios can play a major role in all three aspects of the assessment.

As one surveys the model, it should be remembered that it assumes that the only appropriate focus for any of the three levels/

types of assessment is gain or change over time. That means that the first step at any level is to produce baseline data at some initiatory point (beginning of year, beginning of school, initiation of a program) against which performance at other points in time can be measured. Consideration of that proposition leads quickly to an understanding that obtaining baseline data about student performance is critical to the whole assessment model. Portfolios offer one means of capturing baseline data and adding data over time which can clearly show gain or change.

Figure 2

ASSESSMENT IN PERFORMANCE MANAGEMENT



Student Assessment

As shown in Figure 2, student outcomes should be defined more broadly than scores produced on achievement tests. In many cases, student attitudes and self-management behaviors must be changed before academic performance can improve. In some cases (e.g., severely handicapped students, preschool age children), attitudes and

behaviors are the teaching-learning focus rather than academic content. Therefore, measures of student outcomes should encompass academic outcomes, attitudinal changes and something which might be labeled intellectual growth.

In discussing the uses of portfolios in student assessment, Howard Gardner suggests that they lend themselves to assessment of products, perceptions and reflections. In that framework, measurement of academic outcomes might be seen as assessment of what the student is able to produce.

When measuring changes in attitudes, we are usually measuring changes in students' perceptions and feelings, changes which may be very important to what and how they produce. Gardner, Wolf, and others involved in portfolio assessment have found that assessments of this type offer rich insights into student perceptions at various points in time, if portfolio designs require the inclusion of materials that can be reviewed for this dimension of performance.

Obviously, I have defined academic outcomes and intellectual growth differently. It may not be a very valid separation, but the term "intellectual growth" is used here to try to identify the potential of what Gardner has called student "reflections." What and how students thinks about their own work, progress, growth, development is the focus. One might talk about this area as thinking skills, but thinking skills are essential parts of the other two areas identified for assessment as well. Foremost, the separation of this area from the others is meant to suggest that student reflections will not be forthcoming unless they are designed into the assessment methodology.

Personnel Assessment

Personnel (teacher, administrator) assessment should focus on student outcomes, but if one does not know what inputs produced the outcomes, there is little chance of improving outcomes, especially school-wide outcomes. Therefore, what the professional educator knows and is able to do (competence), his/her application of effective teaching or administrative practices (i.e., practices proven to produce higher outcomes), and the satisfaction of those for whom he/she is responsible (an important ingredient in classroom and school climate) are also important focuses of assessment.

Instructional practices are evidenced in portfolio collections. Commonality of student approaches to problem solving result from teaching not inspiration. Systematic errors in written work across a class of students or a school reflect instruction. Portfolio entries made by teachers and comments by teachers on student entries provide insight into instructional values. The types of tasks and

projects included in portfolios speak of instructional methodologies as well as curriculum content. Hiebert and Calfee[16] suggest that student portfolios provide links between instruction at several grade levels. If the linkage is there, analysis of portfolios at several grade levels should demonstrate it.

Program Assessment

When one changes the assessment lens to focus on the quality and success of a program or school, portfolios also contribute in a variety of ways. If we think of programs in terms of inputs, processes, and outcomes, it becomes easier to see where and how these contributions are made.

Program inputs are usually defined as goals and objectives, characteristics of the target population, available resources (human and fiscal), facilities, organizational structure, and other such variables.

As we consider inputs, Grant Wiggins[14] makes an interesting statement about schools and standards:

> A school has standards when it has high and consistent expectations of all learners in all courses. High standards, whether in people or institutions, are revealed through reliability, integrity, self-discipline, passion and craftsmanship.
>
> Alas, it is thus not too strong to say that many schools exhibit no standards.

If objectives/specifications/standards/exemplars/desired outcomes are clearly defined for student portfolios, much can be learned about the program expectations and goals. Are the standards/exemplars high? Are they short-term? Longitudinal? Have exit standards been established? The absence of these elements also tells us much.

Study of portfolio specifications and guidelines also tells us a great deal about target audiences. Wiggins[14] argues,

> If we are to obtain better quality from schools, we are going to have to challenge the current low expectations for all students in a course, age-cohort, and entire school population.

Are the portfolio standards/exemplars for all students? Are there differentiated standards? How much variance in performance will be allowed? For whom? Initial portfolio plans (they may change over time) also contribute information about school and program organization. Who can enter materials or comments? Who contributes to assessment? What kinds of entries can be included? Responses to

these issues provide insight into faculty and subject matter organization and student involvement.

When assessing the results and impact of programs, it is at the input stage that extensive data regarding pre-conditions should be collected. Since most portfolio designs call for the collection of student work samples at the beginning of the year or portfolio initiation, analysis of the quality of these samples across portfolios offers some information about the state of curriculum, instruction, and learning prior to program initiation as well as a baseline against which to measure the progress made by all students over time.

In program assessment, processes include elements such as curriculum, instructional practices, parent, and community involvement and professional development of educators.

Portfolio contents, when viewed collectively, give great insight into curriculum emphases. There is tangible evidence of subject matter knowledge learned and/or emphasis on communication skills or thinking skills or artistic skills or problem-solving or whatever other emphases have been consciously or unconsciously stressed. Where individual differences in student learning styles or interests or ability have been consciously addressed by program staff, a survey of student portfolios should confirm that. When conscious attempts have been made in a school or program to integrate disciplines and subject areas, portfolio entries can provide evidence of the results.

Lorrie Shepard[17] argues that better student assessments are needed because current tests narrow the content taught. In other words, curriculum tends to focus on what is tested. She also argues that the content of all assessments must be negotiated at some level or another. Therefore, what appears in an assessment represents some kind of consensus building process regarding curriculum. If portfolios are being used as student assessment devices, a survey of their contents should indicate whether curriculum content is narrowing or expanding. Further, portfolio specifications, guidelines, and contents should alert the program assessor to the levels and types of curriculum consensus that have been or are being built.

The contributions of student portfolios to personnel assessment have already been discussed. We can simply reinforce here the notion that at the program level the emphasis should be on instruction not individual instructors. To determine the quality of and emphases in instruction, we must look across portfolios not within a single portfolio.

Student portfolios may or may not provide information about school environment or parent involvement. It depends upon the types of information collected and placed in the portfolio. For ex-

ample, several of the Next Century Schools projects are collecting periodic assessments of student self-concepts. However, these may not become part of a student's portfolio.

Student portfolio contents may not give much insight into the professional development of teachers and administrators, but the presence and design of the portfolios can. Hiebert and Calfee[16] conclude that "student portfolios provide vivid and engaging content for professional discussion and collegial sharing." The conclusion is supported by this author's experience. Successful portfolio assessment projects that were designed without this dialogue and sharing appear to be non-existent. Successful projects, in which professional dialogue among program/school staff about the quality and meaning of portfolio entries is lacking, also appear to be very infrequent, if not non-existent. The presence of student portfolios offers the program assessor several avenues for dialogue with administrators and teachers about the professional growth and development that is taking place.

Since the primary function of student portfolios is assessment, their presence and contents should provide the program assessor with direct information about the alignment of curriculum goals, instructional strategies, and assessment activities. Collecting information about these alignments has long been an issue and intent of curriculum evaluation.

Program outcomes are inclusive of student academic achievement, affective development, attitudes and behavior, teacher and administrator morale, and changed school/program organization. When assessing program outcomes, student portfolios should contribute greatly. If clear performance standards/assessment targets have been created, individual and collective student achievement against those standards can be readily measured. Progress of students of different types and levels should be easily identifiable.

The "biography of a work" which Wolf describes as a product of portfolio development can, in program assessment, be translated to a biography of students' works in which one can read a number of outcomes. Perhaps one of the most important outcomes at the program level will be the consistency of performance across students. If the challenge to low expectations for all students in an age-cohort, class, or program for which Wiggins argues has been mounted, differences in student performance outcomes should be minimal; i.e., they should be within narrow, tolerable limits.

Student affective and attitudinal changes as well as academic progress can be assessed to some degree in the construction, characteristics, and quality of the work produced over time. If student perceptions and reflections as well as performance are valued and devel-

oped, evidence should exist in portfolio contents. "By looking across portfolios, we begin to see where people excel or flounder," as Wolf[5] contends.

If student portfolios are multi-year endeavors, their contents, specifications, and guidelines are bound to change as the professional staff involved with them change and grow and as the school or program organization changes. Analyses of these changes offer insight into organizational and professional outcomes as well as student outcomes.

It may appear that more attention has been given to the uses of student portfolios in program assessment than to their uses in student assessment. Certainly, more space has been given. However, the approach here was a conscious, purposeful one. Many of the questions about portfolio application to student assessment should be answered in the next section of the paper devoted to portfolio design issues. In addition, my review of several reports of LEP program evaluations indicated that these program assessments were superficial at best. An attempt has been made in these last few paragraphs to suggest ways of collecting and analyzing data from student portfolios which can be of much use in determining the quality and success of an LEP program.

Portfolio Design Issues

Past And Present Problems

The development of portfolios of student work and learning products will be of little value to formal student assessment unless portfolio structure, contents, and evaluations of contents are carefully designed before portfolio development is undertaken. The problems of the past must be resolved.

Historically, attempts to use portfolios in assessment have met with six problems. Expectations (objectives) of those conducting instruction and assessment have been unclear to both evaluatees (in this case, students) and instructors/evaluators. Guidelines for number and type of inclusions have been nebulous or non-existent; thereby, reinforcing evaluatees' beliefs that "if some inclusions are good, more are better." The results are sizeable, uneven, unequal, and sometimes unrelated stacks of materials and products constituting evidential bases for assessment decisions. Procedures for scoring or rating portfolio entries, combining assessment results, and clearly communicating student outcomes to students and parents have not been clearly thought out and communicated to those who need to know. A clear decision about the measurement construct to be used

in analysis of portfolio entries and use of those results has been lacking; i.e., "Are portfolio entries to be used in a criterion-referenced evaluation context (student development over time) or a normative context (comparison of accomplishment among students)?" Entry and analysis procedures have been unclear; i.e., questions such as the following have not been thoroughly discussed and resolved in advance of implementation of the portfolio process:

- Who (students, teachers, others) can enter materials?
- Who (students, teachers, others) participates in assessment? How often? Under what conditions? What standards will be employed?
- For what period of time will portfolio entries be kept? For how long are they valid indicators of progress or accomplishment?
- Who has access to the portfolio and the evaluation results?
- What procedures will be used to delete entries from the portfolio, when and if necessary?

Persons given the task of evaluating portfolio entries have been given little or no training in how to evaluate them and few standards against which to measure progress. The results are high inference and subjectivity.

Portfolio Design Questions

The questions below can form the skeleton of portfolio design. Designers may wish to add others that address uniquenesses in their students or settings.

1. What instructional goals, objectives, and outcomes do we want to measure?
2. Which ones (goals/objectives/outcomes) are not now being assessed adequately by other means?

   NOTE: Don't reinvent the wheel. If current assessment methods are adequate, why switch?

3. Will portfolio entries and their analysis be used to assess individual student progress over time or to compare student accomplishment taking into account individual differences?

NOTE: Will the portfolio be used for criterion-referenced assessment or normative assessment? The answer will dictate much about types of entries and procedures for entry.

4. What evidence of progress and/or accomplishment will be required? What evidence of progress/accomplishment will be allowed?

   NOTE: The first question addresses the need for a consistent base of information from student to student. The second addresses issues of individual differences such as creativity, best effort, learning styles.

5. Who will select entries? Why?

   NOTE: In some plans, teachers select all entries. In others, students build their portfolios within specific guidelines. Several researchers and developers recommend that both parties be contributors. What about administrators? Parents? Obviously, age of students, content area and other factors need consideration.

6. What types of evidence can/will be accommodated in the portfolio? Why?

   NOTE: This question was addressed in an earlier section of the paper where it was stated that the type of portfolio, its storage and retrieval system, the area(s) of content involved in the assessment and the characteristics of the students have to be considered.

7. How will portfolio contents be rated/scored/judged? Used in student valuation? Program evaluation? Instructional improvement?

   NOTE: These questions require resolution of both measurement and evaluation issues.

8. Who (students, teachers, others) will contribute to the assessment?

9. How will assessors be trained? What controls will be used to assure some degree of validity and reliability in assessment results?

10. How will results of portfolio assessments be communicated to students? To parents? To the school district?

11. How and when can/will portfolio entries be deleted?

12. What can/will we learn about the success of our program/project from the analysis of student portfolios?

Issues For Discussion

Underlying the twelve questions above are a number of philosophical and measurement issues that need to be discussed and some agreement reached by the professionals in a program, school, or school district before portfolio utilization is undertaken. Perhaps the discussion is best facilitated by development of propositional statements such as those below which are offered for debate. They are a compilation of many of the premises found in the current literature on student portfolios.

1. Portfolios can best be used to assess a student's ability to produce, perceive, and reflect.

   NOTE: This statement is attributable to Howard Gardner, Harvard University (see references).

2. Portfolio entries should be selected by both students and teachers by mutual agreement. Both parties have a stake in the teaching/learning process.

3. In program assessment, portfolios provide insight into process as well as products and outcomes.

4. Portfolios are best used to assess student development over time rather than to assess comparative accomplishments of students.

5. In the arts and humanities, the versatility of the student should be assessed.

6. Portfolios do little to accommodate learning styles unless students are encouraged to produce and submit diverse types of materials and products.

7. Portfolio development and cooperative learning activities go hand-in-hand. (The two can be easily related.)

8. In areas such as writing, evidences of the whole process are more useful than the final product(s) alone.

9. If student reflection is desired, both self-critiques and teacher critiques of entries are required (so that teachers and students can compare them).

10. Evaluation of portfolio contents requires at least two levels of organization: categorical organization of raw data/evidence and summaries or syntheses of available data.

11. If portfolios are to be used in assigning grades, scale descriptions for the requirements for A, B, C, D etc. must be developed.

At least two additional propositions for debate among those contemplating portfolios for LEP students should be added to the list:

12. Some, but not all, written and oral portfolio entries should be in the student's native language. The choice of which entries will be in English and which in the native language should be the student's.

13. Raters/scorers of portfolio entries by LEP students must include at least one person proficient in the student's native language.

Scoring/Rating Portfolios of LEP Students

Rating or scoring portfolio entries may be as simple as scoring a set of responses to a mathematics quiz in which problems have right or wrong answers; i.e., some entries may be sorted on the basis of right or wrong, accurate or inaccurate. However, that often is not the case. Many entries require the exercise of professional judgment. For example, musical compositions, photographic essays, various pieces of writing, and videotaped performances require more of the rater and rating system than has been typical in many testing programs. The issues are compounded when limited language proficiency and/or the use of multiple languages are added to the situation. At least six elements are needed to properly conduct the scoring/rating process.

As indicated in earlier comments, standards for performance must be predetermined when rating portfolio components. What constitutes an outstanding performance? An acceptable performance? An "A"? The standards should be, as Wiggins[14] suggests, exit standards; i.e., they must be standards that describe acceptable performance at the end of the educational process. In some cases, those standards will be end-of-the-year standards. In others, they may be school exit standards. Acceptable progress toward those exit standards should be judged in terms of movement along a continuum from each student's entry point to the exit standard. Portfolio entries of all students should be judged against prescribed standards, not against each other. For LEP students, the performance standards should address language standards as well as other elements.

Both students and teachers need exemplars of performance at the prescribed standard. What does an outstanding musical composition look and sound like? An acceptable short story? An award winning photographic essay? If native language or mode of thought is to be used, exemplars in the language need to be provided.

Many products and performances probably will be rated on a scale of some sort. This tends to be true even when numbers of points are awarded for the presence of certain features in the product/performance. Usually, the points are totaled and applied to some pre-determined scale. (Our typical grade structures operate like this.) Scales used in rating entries should be behaviorally anchored scales; i.e., each point on the scale should be described in terms of the behaviors required to achieve that level. What elements of performance must be present to achieve a "5" (on a five-point scale)? What elements can be absent and still allow the producer to obtain a "3"? If exemplars of exit level performance have been provided (e.g., writing performance at the end of the high school years), what elements must be present to obtain an outstanding ("5") rating at the end of the middle school years?

A fourth element necessary to rating and scoring is the use of multiple raters/evaluators. Olympic competitions rely on multiple judges. If portfolios are to be a serious part of student assessment, an approach not unlike that used to score Advanced Placement Examinations should be used. A team of raters (at least two) will add validity and reliability to the assessment score. Further, the use of multiple raters is essential in assessing portfolios of LEP students. If native language is allowed, one or more members of the rating team will need to be proficient in the native language. If entries make use of only the English language, there is still need for at least one rater to be proficient in the native language. He or she will be the person more likely to identify the characteristics of the product or presentation directly attributable to language and bring these to the attention of colleagues.

Although it may not always be essential, this writer recommends the use of consensus processes among raters. Rather than supplying two or three independent ratings/scores which are then averaged, each rater generates independent ratings, then meets with colleagues. Ratings and rationales are shared, and the group arrives at a consensus rating and a consensus rationale for that rating. While this approach requires additional time, it strengthens validity and reliability of the final scores, contributes to the comfort and "assessment literacy" of raters, and provides staff development both in assessment and instruction. Rater teams always seem to talk about what can be done to improve student performance.

If the reader has followed closely the five rating/scoring process elements described thus far, he/she can predict the sixth. Raters/scorers must be trained. They must be trained in how to apply the standards, exemplars and rating scales to the student products. If consensus is to be used, they must be trained to use the consensus process.

A Final Comment

New student assessment technologies, including portfolios, can provide new and often better information about student performance and development and about program performance than has previously been available. There appears to be great potential in the use of portfolios with LEP students. However, the value is yet to be determined. Experimentation, perhaps as much as ten years of it, will be needed. Thankfully, that experimentation is underway.

Notes

[1] National Governor's Association. (1991). From rhetoric to action. Washington, D.C.: Author.

[2] National Business Roundtable. (1990). Essential components of a successful education system. Washington, DC: Author.

[3] National Center On Education And The Economy. (1990). Seven components of the restructuring process. Washington, DC: Author.

[4] French, R. L. Teaching strategies and learning processes. Tennessee Education.

[5] Mehrabian, A. (1971). Silent messages. Wadsworth Publishing Company.

[6] Hall, E. T. (1969). The hidden dimension. Anchor Books. Additional information can be found in Hall's works, The Silent Language and Beyond Culture.

[7] Ascher, C. (1990). Assessing bilingual students for placement and instruction. ERIC/CUE Digest, (No. 65). New York: Clearinghouse On Urban Education.

[8] Wolf, D. P. (1988, December, January). Opening up assessment. Educational Leadership, 45(4), 24-29.

[9] Paulson, F. L., Paulson, P. R., & Meyer, C. A. (1991, February). What makes a portfolio a portfolio? Educational Leadership, 48(5), 60-63.

[10] Brandt, R. (1990, March). On assessment in the arts: A conversation with Howard Gardner. Educational Leadership, 47(6), 24-29.

[11] Wolf, D. P. (1989, April). Portfolio assessment: Sampling student work. Educational Leadership, 46(7), 34-40.

[12] Resnick, L., & Associates. (1991). The New Standards Project: An Overview. University of Pittsburgh, Learning Research and Development Center.

[13] Stiggins, R. J. (1991, March). Assessment literacy. Phi Delta Kappan, 72(7), 534-539.

[14] Wiggins, G. (1991, February). Standards, not standardization: evoking quality student work. Educational Leadership, 48(5), 18-25.

[15] For more information about Linda Vista Elementary School and its electronic portfolio, contact Dr. Adele Nadeau, Principal, 2772 Ulric Street, San Diego, California 92111.

[16] Hiebert, E. H., & Calfee, R. C. (1989, April). Advancing academic literacy through teachers' assessments. Educational Leadership, 46(7), 50-54.

[17] Shepard, L. (1989, April). Why we need better assessments. Educational Leadership, 46(7), 4-9.

Response to Russell French's Presentation

Alice J. Kawakami
Pacific Regional Educational Laboratory, Hawaii

I would like to thank OBEMLA and OERI for inviting me here as a discussant. Although most of my experience in education in Hawaii has not been with bilingual education, Hawaii's public school population is multi-ethnic and students come from various language and dialect backgrounds. I have been involved in developing a language arts program which has been attempting to develop portfolios as a means of assessing elementary students' learning. With this background, I find Dr. Russell French's paper insightful and important in focusing on some of the key elements of authentic assessment within a classroom setting.

As I consider the theme of this conference, "Achievement as a Child's Universal Language," it seems that one way for our students to participate in that universal language and to be able to speak of their achievements is to provide them with opportunities to give voice to their successes through the use of portfolio assessment. As Dr. Russell French points out, current standardized tests do not adequately "showcase" the learning of our students. I will speak today on my experiences with the Kamehameha Elementary Education Program (KEEP) in Hawaii, during the development of portfolios. KEEP is a language arts program for elementary students. It was developed to assist Native Hawaiian students in the public schools. In its early years, the program was grounded in culturally congruent interaction styles, classroom organization, comprehension focused direct instruction, and a mastery learning system to track student progress. Recently, the curriculum was expanded to maintain the cultural component and include the development of writing and, most importantly, to focus on the development of students' ownership of their learning. This required a paradigm shift in curriculum and assessment as well as on the part of the teachers. The relevance of that work to our topic today is the process of moving from assessments which were standardized to authentic assessment. The focus of my comments will be the support needed for teachers undertaking this change in teaching and learning and the documentation of more functional, authentic learning.

With this change, instruction and assessment took on a different, more responsive face. The transition from a reliance on test results to portfolio assessment is still in progress. Many of the issues raised in Dr. French's paper were addressed in the process of program development at Kamehameha. I would like to describe some of the major changes and accompanying support that was needed. Moving to

portfolio assessment was not easy because it called for a reexamination of our basic beliefs about the role of teachers and students and the criteria by which we measure success.

In retrospect, we found the framework of a paradigm shift helpful in understanding the changes we have made. The framework enabled us to look at some of the assumptions underlying the original program (based on a transmission model of instruction) and the current whole literacy program (based on a constructivist model of instruction). The following table outlines four areas of change. The content and process of instruction refer to the actual classroom routines. Assessment and evaluation refer to the monitoring system that directs the focus of instruction. Accompanying these routines in the classroom are perceptions about the roles of the teacher and the student.

A Paradigm Shift in Teaching and Learning

	Transmission Model	Constructivist Model
Instructional Content/Process	predetermined classroom content	constantly constructed content within class
Assessment	external criteria	internally developed criteria
Role of teacher	giver of knowledge	facilitator of learning
Role of student	receiver of knowledge	coordinator of own learning

Under a transmission model, content and instructional strategies are fixed by curriculum guides and published materials. In a constructivist model, content is based on curriculum areas but negotiated with student input on topics of interest. Assessment in the transmission model is usually dependent upon externally set criteria such as skill-based mastery tests. Assessment in the constructivist model is based on goals set collaboratively by the teacher and student as the criteria for success. This criteria arises from the context of the classroom and is tied to benchmarks for student progress. It provides feedback for learning, and is useful to students, teachers, and parents. This is the critical function that portfolio assessment

provides in the KEEP whole literacy classroom. It is responsive to the classroom learning environment and should be authentic and meaningful. These models of teaching and learning are supported by the roles of teacher and student as indicated in the table.

In order for teachers to support student learning through the use of portfolio assessment, a shift must occur in their perception of their role as teacher. This shift may be difficult for many teachers because it is predicated upon change in the assumptions about teaching and learning. Support for making this shift is critical because it calls for major philosophical change. All change is difficult to bring about and this is not an exception. Our experiences at KEEP taught us that these changes cannot be mandated but need to be developed. The development is dependent upon the extent to which staff development activities can be grounded in the same constructivist model.

The following change matrix is helpful in understanding the change process and the kinds of support needed by teachers. These changes must be viewed as part of the process of developing portfolio assessment. With tests as assessment, the guidelines for administering the tests were made explicit in the testing procedures. With portfolio assessment, guidelines for developing a portfolios are grounded in the teacher's attitudes, beliefs, and values which require collaboration with students. The curriculum, classroom interactions, teacher development, and student development are an integral part of the implementation of portfolios.

In moving toward a classroom environment where the teacher acts as a facilitator of student learning and the students take responsibility for their progress, teachers move from a relative position of isolation and reliance on curriculum materials and guides to a context of professionalism and collegiality. The following change matrix is useful in understanding the steps involved in moving from one point to another within the framework. Through our often frustrated efforts, we found that institutional support for teacher development along these lines is critical.

Change Matrix

Content	Context		
	Isolation	Congenial group	Professional colleagues
Materials	A		
Behavior/ Strategies		B	
Attitudes, Belief, Values			C

We can view the process of change on three points in this matrix. Initially, at point A, teachers may find themselves operating in the isolation of their own classrooms. Their primary guides for professional development are the instructional materials used in their classroom. This is ideal for teachers operating under a transmission model, with fixed scope and sequence charts, lesson plans and assessment instruments developed by publishing companies far from the contexts of their classroom. As a teacher begins to seek the input from colleagues, point B may describe the interactions. Here, teachers meet in groups and discuss students' behaviors and discuss strategies they have tried. This stage lends itself to more dynamic views of teaching, and more experimentation based on classroom conditions. A teacher who is operating in a constructivist model would be placed at point C, Here, attitudes and beliefs about teaching and learning, strategies for teaching and instructional materials are discussed with colleagues. Teachers at this stage reflect upon their teaching, seek feedback from others, and take responsibility for their own professional development. This stage is analogous to students who have assumed responsibility for their own learning.

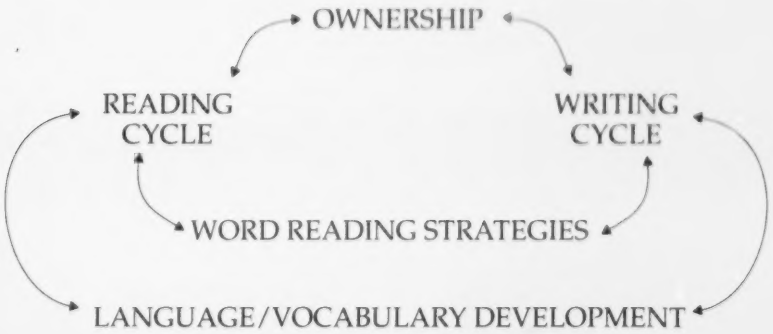
When we first began our training, we thought that by giving teachers whole language instructional materials, we would instantly make them constructivist teachers. We created opportunities for groups of teachers to meet and discuss materials and that is exactly what we obtained, groups of teachers looking at big books. Our next step was to provide them with training in teaching strategies. We provided workshops on writing process approach and shared readings. Still, change was negligible for the majority of teachers. We finally realized that to effect long-term, deep change in teachers, we needed to address their attitudes and beliefs. We took the approach

of focusing on those underlying assumptions in day-long retreats. We posed questions for discussion about the learning process that compelled teachers to examine their own learning process, their beliefs about the best learning environment for their students, and the most appropriate role for them to take vis a vis their students. At this point, we began to see progress in developing reflective, responsive teachers who began gaining confidence in their abilities to function as facilitators of student learning. Until that moment, the need for portfolio assessment was not realized.

In our discussions of shifting the focus of student learning goals, we had to develop an innovative means of conveying the objectives of the language arts program in a way that was not tied to a format with the limitations inherent in a scope and sequence chart. We decided to begin by presenting teachers with a diagram of aspects of literacy to be developed within the program. There are six aspects of literacy in the KEEP curriculum. The following diagram shows the presentation format that we used.

KAMEHAMEHA ELEMENTARY EDUCATION PROGRAM

SIX ASPECTS OF LITERACY



VOLUNTARY LITERACY
(Reading/writing out of school)

The first and most important aspect of literacy is the development of students' ownership of their own literacy learning. This

translates into activities which operate within meaningful contexts. The purpose of reading and writing activities in school is to communicate ideas that are relevant to the students rather than to complete a number of worksheets or to move through a specified number of pages in a practice book. This provides authentic literacy activities. If the content of the communication is meaningful to the students, instruction in word reading strategies and language and vocabulary development is purposeful as a critical part of communicating. If reading and writing in school are meaningful, the application of these abilities should flow beyond the realm of classroom work and into the realm of voluntary reading and writing. This voluntary literacy, outside of the school, is the application of school learning and a demonstration of students' ownership of learning. This concept of the curriculum as the basis for real application of school learning is the perfect context for the development and use of portfolios. It is also an absurd context for assessing learning through standardized tests.

Teachers who are committed to developing students' ownership of functional meaningful literacy recognize the artificiality of testing developed under a transmission model. However, until teachers are committed to this concept of literacy development and have recognized the attitudes and beliefs that underlie their own teaching, portfolios may be just another requirement of the curriculum. When teachers sit with their students and collaboratively document the kind of literacy learning that has taken place, portfolios take on the role of measuring the complex, functional learning that is needed for success in society. Evidence to document progress is far different from information on skill-based mastery tests. More appropriate documentation would be described in Dr. Russell French's categories of performance tasks, exhibitions, and portfolios. In KEEP classes, projects, reading logs, samples of a student's best writing, and correspondence between students, teachers, and parents qualify as legitimate assessment information. Benchmarks have been developed as ties to state performance expectations.

If bilingual students and students from multicultural homes are to meet success in our schools, assessment must be designed to allow for authentic learning to be a legitimate part of the definition of success. My comments reflect a perspective that portfolio assessment must be negotiated by teachers in classrooms and, for that to occur, teachers will need a lot of support in developing the attitudes, values, and beliefs that free them from the confines of standardized tests. All of the specific issues raised in Dr. Russell French's paper need to be addressed as portfolio assessment is implemented in schools. In addition to those concerns addressing authentic assessment of complex learning of students, there must be attention to the complex learning that teachers will experience as the face of assess-

ment changes to reflect complex, functional learning. The problems that are identified in Dr. Russell French's paper:

> unclear expectations, nebulous or non-existent guidelines, unclear scoring procedures, lack of definition of the measurement constructs for portfolio entries, vague entry and analysis procedures, little or no training in implementing portfolios, and few standards were all a part of our process to shift to authentic assessment. Although all of these issues have not been completely resolved, teachers must be involved in the development of portfolios, because they will be at the delivery point of this new assessment format.

The technical questions raised in Dr. Russell French's presentation are vital to the development of authentic assessment. My remarks were intended to point to the equal importance of the teacher and the support that is needed to implement changes in assessment. Someday, our students will be able to use portfolio assessments to support their success in learning and allow us to recognize their real achievements.

Response to Russell French's Presentation

Daniel Koretz
RAND Corporation

A lot of the issues that have been raised so far pertain to portfolios in general not specifically. Their use is with children who are limited in their use of the English language. I'm going to skip over a lot of the issues that are more generic, but I do need to touch on a few to make my more specific comments clear. I should clarify at the onset that I'm speaking as a proponent of portfolio assessment. I've been involved in developing the portfolio assessment program for more than three years. I do think that, properly done, portfolios have a substantial potential for improving instruction, but I think that they're very difficult to do and, if we don't go into our efforts to use them with eyes open, we stand to lose more than we gain. Moreover, I think that it's very doubtful that performance assessments will provide the kind of opportunity that was alluded to twice in the past two talks, that is revealing abilities of LEP children that have not been revealed by traditional tests. I think they may represent a real opportunity to LEP children, just as they represent a real opportunity for any children, and that they may help steer teachers, as Dr. Alice Kawakami suggested, toward more interesting, engaging, and demanding course work. This pertains to all students, regardless of their native language. I'll come back to why I'm a little more skeptical about their usefulness in revealing abilities of LEP children that have been hidden because of their difficulties with English.

A few generic comments. First, portfolio assessment and performance assessment, in general, really have two different goals. The portfolio's assessment goal is one of getting better assessments of what children can do. Better in the sense of tapping abilities or capabilities that traditional tests might not, and performance assessment is improving instruction. Those two goals are very different despite the fact that proponents of performance assessment often talk of them both at the same time and often assume either implicitly or explicitly that, if a task or an assessment is authentic, and I'll leave open what that means, it will improve instruction and provide good assessment. I think that's simply not true in many cases. It's very easy to come up with tasks that are engaging in the classroom and potentially very useful in the classroom but have no discernable measurement value. There are some cases where they do overlap, but I think the overlap can easily be overstated. I think the conflict between those two goals or the uneasy compromise between those two goals is particularly severe when the children are being tested at limited proficiency in English, and I will come back to why that is so.

As a second generic caution, portfolio assessment is extremely difficult. It is not hard to come up with a collection of work from an individual student that the teacher and the student may agree has been properly evaluated. It is extremely hard to get collections of work from large numbers of students that are rated in a way that is even halfway comparable from teacher to teacher, school to school, child to child, which is what large scale assessment programs have to do.

I've been involved in the Vermont performance assessment work since it started more than three years ago. Most of the participants remain very enthusiastic about portfolios, very optimistic that they will, in fact, help improve instruction, but the list of difficulties that the participants have faced in the past three years is very long. I will just list a few of them for you. One is that because portfolios can include, as Dr. Russell French mentioned, a wide and diverse array of materials, raters often find that they get work in portfolios that they can't rate. Once they have agreed on standards, on criteria for judging student work, lo and behold, children produce things that don't fit.

The converse of this is raters who report that they periodically, in fact, not too infrequently, come across good work that they recognize as hard work from a capable person that slips through the cracks, because it was not the kind of work that was in mind when people designed the criteria. Raters reported that it was extremely difficult to aggregate the ratings of individual tasks because a portfolio is a collection of tasks and products submitted to some summary judgment of individual work. A particularly severe problem is that the nature of classroom assignments was often too poorly documented for the raters to judge performance. For example, if a student in a mathematics portfolio does not show adequate explanation of why he or she solved the problem in a certain way or how he or she came up with the answer, is that because the student can't do it, didn't do it, or because the teacher's assignment didn't give the student any reason to do it. Well, often, you can't tell. Let me put it differently. It is very hard to tell, and it requires a lot of careful work to make sure that the relevant documentation is there.

Sometimes raters found that they lacked enough information about students to judge what they were doing. For example, again in mathematics, a given solution to a specific problem for some students might be a remarkable act of invention if that student has never confronted that particular kind of solution before. But, if in fact, another student has worked on that kind of problem at great length and happens to know that one of the ways to solve a problem of this type is to use the such and such method and just regurgitates it back, that is, in a sense, a much lower or at least different kind of performance. How do you know which is which? I mention these not

to discourage you but to encourage you not to see portfolio assessment or performance assessment, in general, as some kind of panacea. It's damned hard work, and it often doesn't work.

Now, one of the consequences of these problems is that in Vermont, in any case, participants are gradually moving toward a view that I have held all along which is that the contents of the portfolio, as Dr. French mentioned, have to be very carefully circumscribed. You have to say to people, "Here are the kinds of things that we can rate, given our criteria, our standards. Here are the kinds of things we either don't want to rate or can't rate." And the reason I want to stress that is because one of the kinds of products that raters in Vermont had difficulty with is non-verbal products. What do you do with a video tape? Well, depending on your criteria and your standards, your exemplars that you give children, a video tape may be perfectly appropriate, but for other sets of criteria, it's unusable, and if, for instances, what you are interested in is the ability to communicate mathematics, a non-verbal video often won't tell you much. Now, what this is leading up to, I think you all can see, is that whether or not a portfolio system is a better, truer gage of what LEP children can do, despite their limited proficiency in English, depends on what you say should go into a portfolio and what you say should not go into a portfolio depends entirely on what you want to say at the other end, what inferences you want to draw about it, about student performance. This is a big open question really right now. I don't think there is firm evidence that portfolio systems will, in general, be harder for LEP students, but that is my suspicion, and there is no evidence that, in general, they will prove better, in the sense of revealing more of what they can do despite their difficulties with English.

Now, some observations that are specifically about language and portfolios and LEP children in portfolios: It's very, very difficult to avoid the confounding of language and other skills when you do portfolios because of what goes in them. First of all, in the case of writing, that is obvious, but even in the case of mathematics, if you are going to do more than a traditional test, you want to find out what children can do, what they can explain, how they did things, and almost inevitably, you start drifting into a mix of whatever other things you want to measure and language.

In Vermont, math and writing were both assessed in grades 4 and 8 this past year, and the raters found that to be a very serious problem, even though in Vermont, as many of you probably know, there are virtually no children with limited English proficiency. It's a very homogeneous state, but even so, the raters often felt that children might be rated high in some cases because their math was good and, in other case, because their language was adept.

I want to make one more point about the confounding with language. That's not necessarily bad; it's a question of what it is that you are trying to measure and what it is you are trying to conclude. Many of you may be familiar with the new National Council of Teachers of Mathematics (NCTM) standards for curriculum and evaluation. Those standards are widely accepted, in name at least, in the education community, and they stress communication of language as one of the primary goals of instruction. An assessment that is designed to match the NCTM standards could not be designed to find out what they know but can't communicate. It is designed to find out what they know and can communicate, so there is no way of avoiding language.

I think that, in dealing with this, we are not getting into the controversial and stickier issues and, in dealing with them, I think it is necessary to separate technical from philosophical issues. Whether or not language is confounded with mathematics in a portfolio may be a technical issue. When you want language proficiency to be measured by a portfolio system, it is not a technical issue. It's a philosophical issue, and I'll give you some examples that are kind of silly because they are so extreme.

Many years ago I gave a Wexler Intelligence Test to an Israeli graduate student at Cornell whose English was fabulous. He had been studying in American schools since seventh grade. His father had been a diplomat. He was then a graduate student in sociology, but a few of the tests really threw him for a loop. One of them, called Digit Span, requires the reciting of even longer strings of numerical digits to the student, and the student is supposed to repeat them back. You see how long a span the student can remember, and then the student has to do it backwards, which is harder, and you are doing it against a stopwatch, which is very unpleasant. Well, this fellow really started to stumble. This is peculiar. So, just out of curiosity, since it was not for a formal evaluation but for practice, I switched to Hebrew, which I spoke, at that point, quite well because I used to live on a kibbutz, and immediately, his performance picked up. And, in that case, you would probably want to give that subtest in the person's native language, because you are trying to draw inferences about something that has nothing to do with language. It has to do with the ability to memorize spans and digits, but what if you are trying to measure, for example, a person's ability to write, or person's ability to explain solutions to mathematical problems, there's no choice. It has to be in some language, and what language should it be in? Well, that's a philosophical question.

The application of portfolios to LEP children underscores the difference between instructional and measurement goals. It's true, in general, for portfolio assessment, and again, this is a philosophical question, there might be a case where you prefer that a student's in-

struction be in English to give that student practice or might feel that student is not yet advanced enough in English that the assessment would be fair in English. There might be times where the reverse is true, depending on the situation, what you are trying to infer from children, what the particular children's abilities and goals are.

Finally, I will wrap this up, because we are running a little late, and I think there should be time for some discussion and argument. Using portfolios with LEP children, opposes some really substantial, practical constraints depending on the decisions you make about philosophical questions. Dr. Russell French raised the question of using raters who are fluent in the student's native language. Well, that might or might not be possible in a district such as Houston, where a large share of the population is LEP and almost every LEP kid in the district speaks Spanish as a native language. It is not possible, even remotely possible, in many other schools. In my neighborhood school, which is not even a high minority school, there are native speakers of Spanish, Portuguese, Japanese, Chinese, Swedish, Norwegian, and Hebrew, and there is not, to my knowledge, a single teacher in that school who speaks any of those languages fluently, let alone enough teachers who speak them fluently enough to test the reliability of scoring. The real world for most districts is that portfolios are going to be assessed by people who speak English and not anything else, and that raises very serious questions about how the portfolios ought to be run for LEP children and how they ought to be scored.

What potentials do portfolios have for LEP children? Personally, I think, in one sense, they could be a very big step forward, and here, I am speaking less about a technical view than a personal view. I think there has been a long and unpleasant history in the United States of giving children who have difficulty with school, for whatever reason, whether it be lack of facility in English or whatever, an even more boring diet of course work than regular kids get. This might be a big step away from that. Rather than taking children who have a little difficulty in math or difficulty in tracking the directions in English and saying, "you're going to do even more drill and more practice, until you are bored to tears" you say, "you're going to do some difficult work that actually makes you think and write, just like anybody else." Will these assessments tap abilities that some of these kids have that standardized tests might not? I think that it is very unlikely. I think that a portfolio assessment is inevitably going to, first of all, put more demands on children, especially LEP children, and second, bring to the fore some very difficult, not just technical but philosophical, issues about how LEP children are to be taught. I'll leave it at that.